GeneSpring GX 9

Quick Start Guide

The information in this guide is for use with GeneSpring GX 9

What is GeneSpring GX 9?.....1
Where do I get GeneSpring GX 9?.....1
How do I get started with GeneSpring GX 9?.....1
Is there a demo experiment I can run?.....1
How can I get more Pathways?.....9
Where do I get GSEA compatible Gene Sets?.....10
Where is my data from GeneSpring GX 7?.....10

What is GeneSpring GX 9?

GeneSpring GX is a powerful visualization and analysis solution designed for use with genomic expression data. It allows you to display and analyze large data sets on your personal computer. GeneSpring GX can analyze data from many sources, including academic, biotechnology, and pharmaceutical organizations.

Where do I get GeneSpring GX 9?

You can download a free evaluation copy of GeneSpring GX that will be fully active for 15 days. Go to http://genespring.com and follow the link for "Download FREE Trial".
If you would like to buy a copy of GeneSpring GX, request a quote from the website, call +1-800-227-9770 (Option 6) or contact your Agilent sales representative or distributor.

How do I get started with GeneSpring GX 9?

After installation of GeneSpring GX, you can get started immediately since it comes pre-loaded with a demonstration experiment to show you the functionality of GeneSpring GX. There is a project called "Demo Project" and contains an experiment and results from a study of HeLa cells that were treated with an unknown compound in a cancer treatment pilot study. We encourage you to explore this demonstration project to get to know GeneSpring GX.

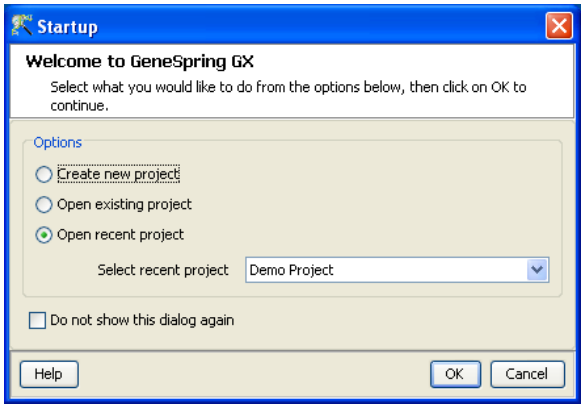
Is there a demo experiment I can run?

You can re-create the experiment from the demonstration project, since the original data you need is provided with the installation. The data files you need are located in the samples directory of the main installation folder, usually found in

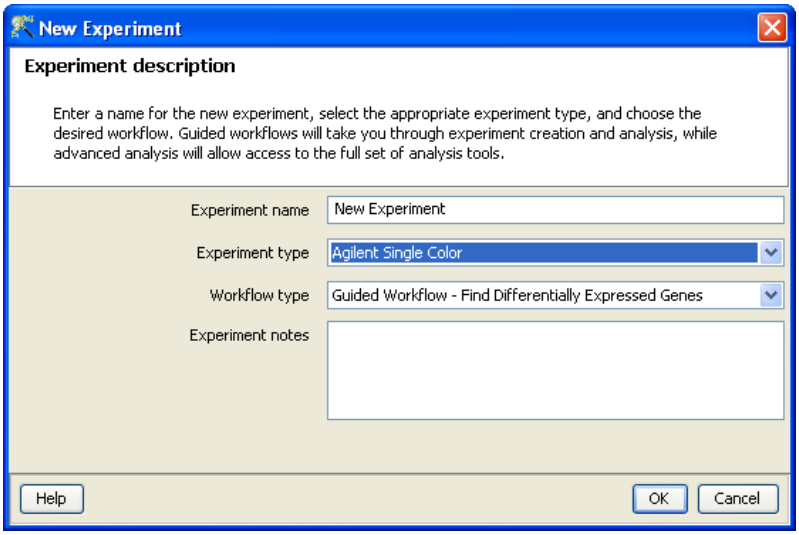
```
C:\Program Files\Agilent\GeneSpringGX
```

Follow the steps below to create a new experiment with the provided data:

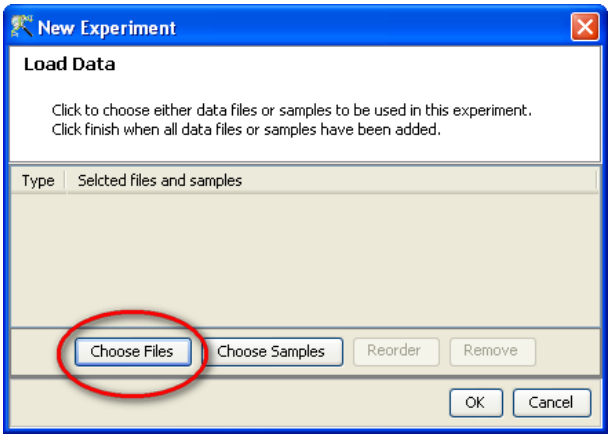
1) Start GeneSpring GX
2) Select the "Demo Project" from the Startup dialog box (or create a new project)



3) This will open the project open the experiment that was stored in this project
4) To create a new project, press the New Experiment icon on the Toolbar ( ) or select "Project -> New Experiment" from the menu bar.
5) This will start the new Experiment wizard that will guide you through the creation of the experiment


New Experiment
Experiment description
Enter a name for the new experiment, select the appropriate experiment type, and choose the desired workflow. Guided workflows will take you through experiment creation and analysis, while advanced analysis will allow access to the full set of analysis tools.
Experiment name
New Experiment
Experiment type
Agilent Single Color
Workflow type
Guided Workflow - Find Differentially Expressed Genes
Experiment notes
Help
OK
Cancel

6) For the name, use something such as "HeLa cells treatment" to give it a sensible name
7) Since the data we provided is a Agilent Single Color experiment, choose "Agilent Single Color" as the Experiment Type.
8) There are two modes of GeneSpring GX, "Guided Workflow" or "Advanced Analysis". The Guided Workflow option will provide a you with full guidance and is a good starting point. Select "Guided Workflow" for the Workflow Type and press OK.
9) The next window will allow you to import new data files into the system or use data that had been previously loaded (samples). For this demonstration experiment we will use new data files.

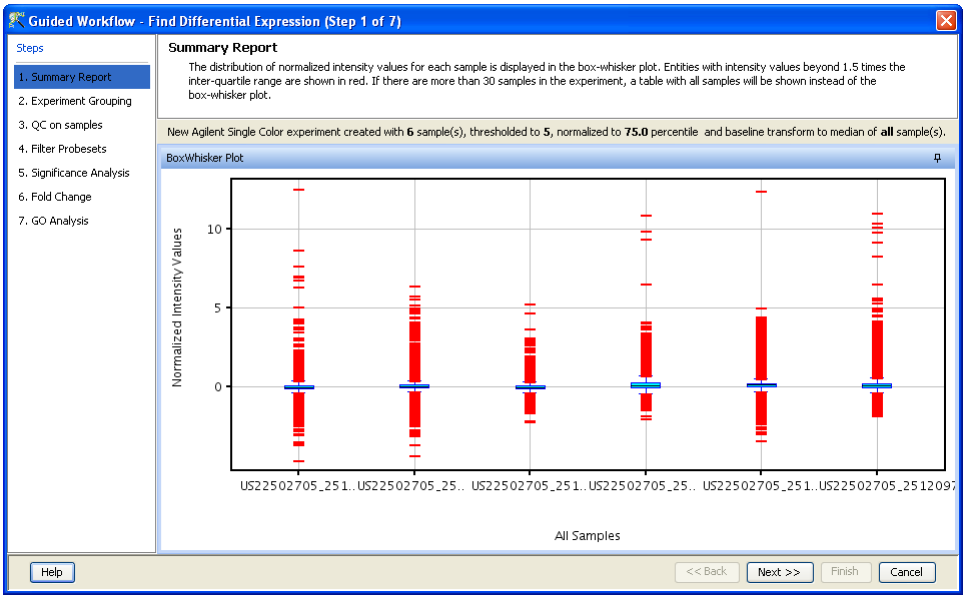


10) Select Choose Files to bring up the Choose file dialog. The choose file dialog will most likely already point to the samples directory in the main GeneSpring GX installation directory, but if not, point it to the directory called "Agilent One-color Demo", which is in

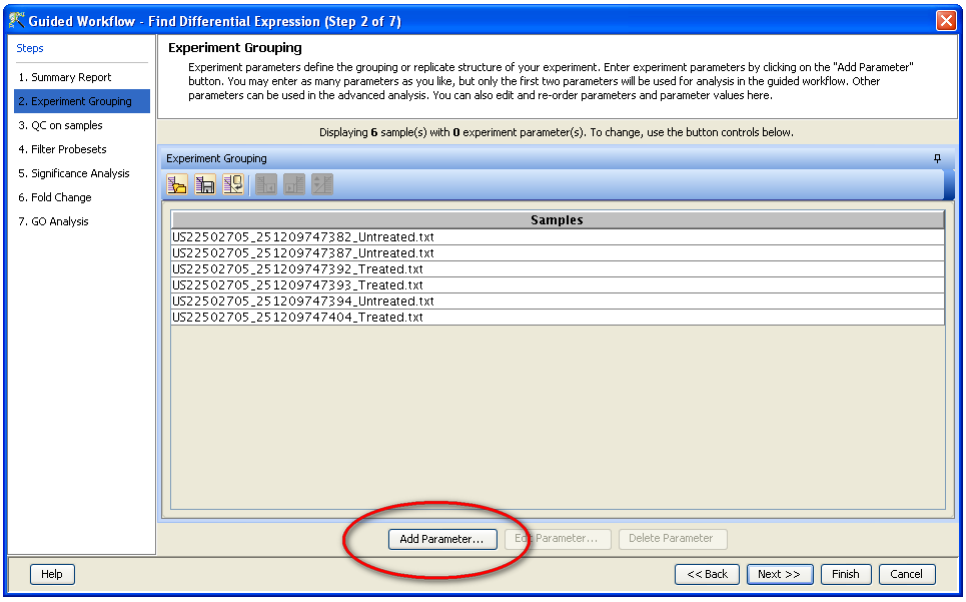
```
C:\Program Files\Agilent\GeneSpringGX\samples\Agilent One-color Demo
```

By default.

11) Select all .txt files to be loaded and press OK:
    - "US22502705_251209747382_Untreated.txt"
    - "US22502705_251209747387_Untreated.txt"
    - "US22502705_251209747392_Treated.txt"
    - "US22502705_251209747393_Treated.txt"
    - "US22502705_251209747394_Untreated.txt"
    - "US22502705_251209747404_Treated.txt"
12) The application will load the files, extract the expression data and perform some standard normalizations and transformations. For the Guided Workflow: Find Significant Differentially Expressed Genes" the defaults are set as follows:
    - Threshold all values to 5
    - Perform a Median Shift normalization to the 75 percentile
    - Perform a Baseline transformation using the median of all samples
13) These are sensible normalization for many experiments and will provide a good starting point. In the advanced analysis workflows you can change these normalization and transformation settings to more appropriate settings for your experiment.
14) After the samples have been loaded and processed, the Guided Workflow wizard will appear and show a summary of the data with a box plot of the data distribution of the 6 samples and an outline of the steps that will be performed on the left side.

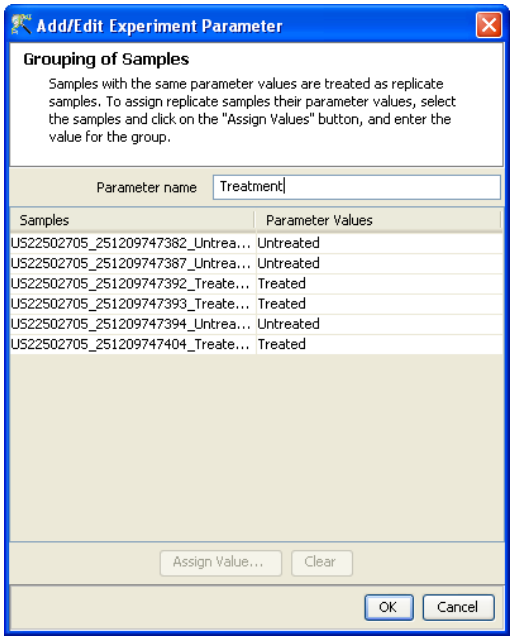

Guided Workflow - Find Differential Expression (Step 1 of 7)
Steps
1. Summary Report
2. Experiment Grouping
3. QC on samples
4. Filter Probesets
5. Significance Analysis
6. Fold Change
7. GO Analysis
Summary Report
The distribution of normalized intensity values for each sample is displayed in the box-whisker plot. Entities with intensity values beyond 1.5 times the inter-quartile range are shown in red. If there are more than 30 samples in the experiment, a table with all samples will be shown instead of the box-whisker plot.
New Agilent Single Color experiment created with 6 sample(s), thresholded to 5, normalized to 75.0 percentile and baseline transform to median of all sample(s).
BoxWhisker Plot
Normalized Intensity Values
10
5
0
US22502705_251..US22502705_25..
All Samples
Help
<< Back
Next >>
Finish
Cancel

15) The guided workflows are controlled by pressing the <Next> button to go to the next step. By pressing the <Back> button you can go back and make some modifications. Press the <Next> button to continue
16) The next window is where we let the program know which samples are the treated samples and which as the untreated samples. This is done by adding an "Experimental Parameter" and pressing the <Add Parameter> button.

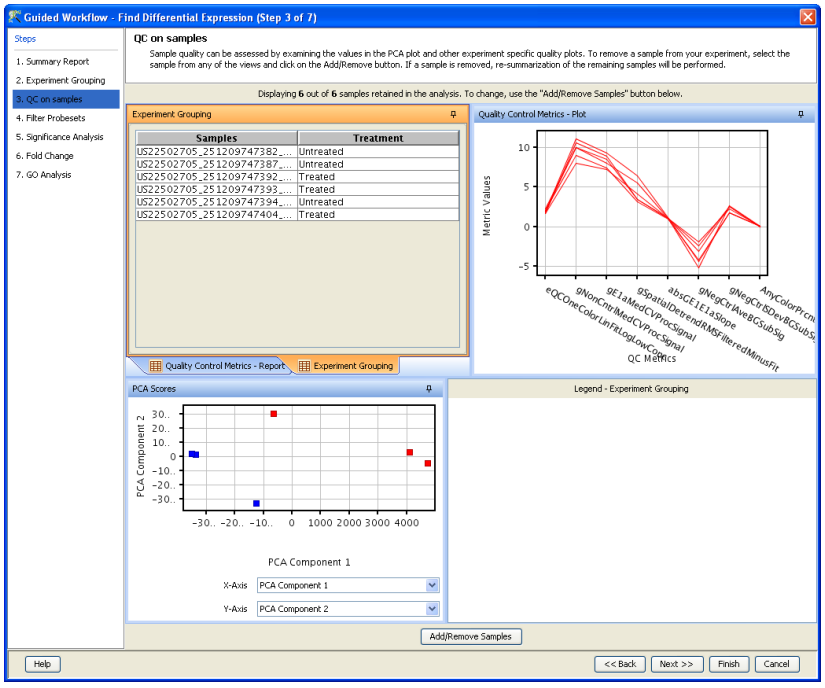

Guided Workflow - Find Differential Expression (Step 2 of 7)
Steps
1. Summary Report
2. Experiment Grouping
3. QC on samples
4. Filter Probesets
5. Significance Analysis
6. Fold Change
7. GO Analysis
Experiment Grouping
Experiment parameters define the grouping or replicate structure of your experiment. Enter experiment parameters by clicking on the "Add Parameter" button. You may enter as many parameters as you like, but only the first two parameters will be used for analysis in the guided workflow. Other parameters can be used in the advanced analysis. You can also edit and re-order parameters and parameter values here.
Displaying 6 sample(s) with 0 experiment parameter(s). To change, use the button controls below.
Experiment Grouping
Samples
US22502705_251209747382_Untreated.txt
US22502705_251209747387_Untreated.txt
US22502705_251209747392_Treated.txt
US22502705_251209747393_Treated.txt
US22502705_251209747394_Untreated.txt
US22502705_251209747404_Treated.txt
Add Parameter...
Edit Parameter...
Delete Parameter
Help
<< Back
Next >>
Finish
Cancel

17) The filenames are pretty descriptive so it is easy to figure out which samples are treated and which are untreated. In the Add Parameter window, make "Treatment" the parameter name and type the values "Untreated" and "Treated" in the Parameter Values columns. (Hint: Select the three samples that need the

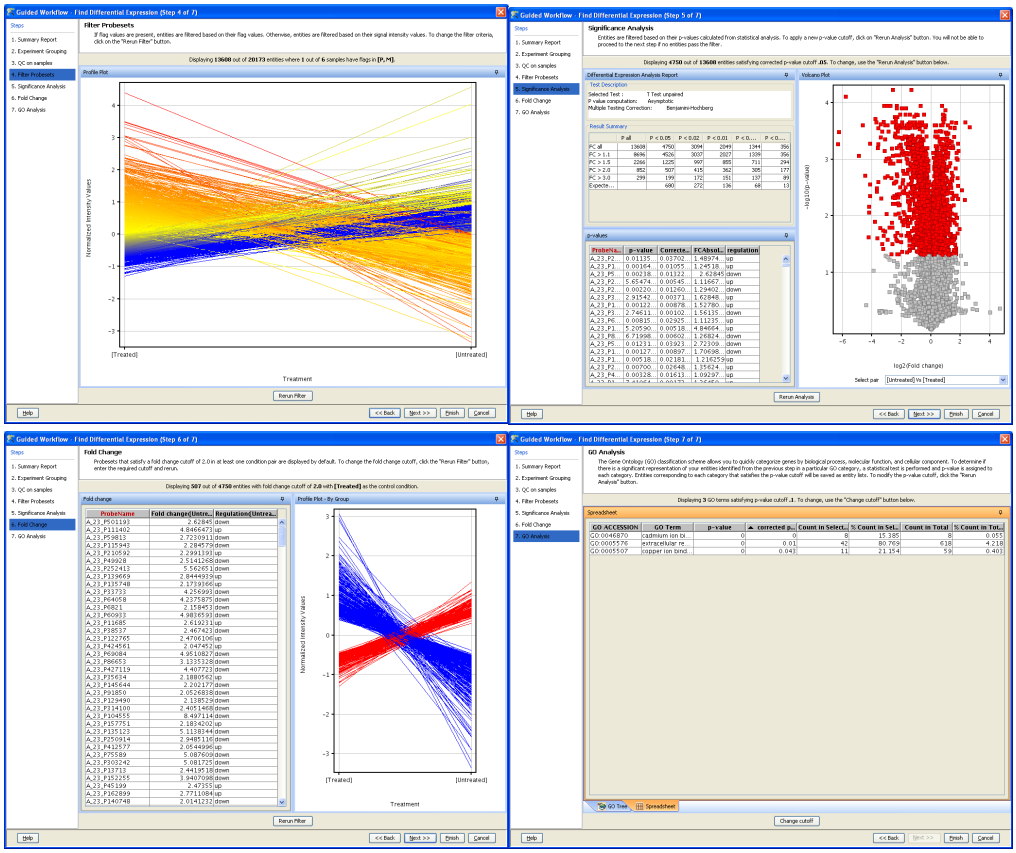
value "Untreated" and press the <Assign Value> to speed up the assignment of multiple values at once)



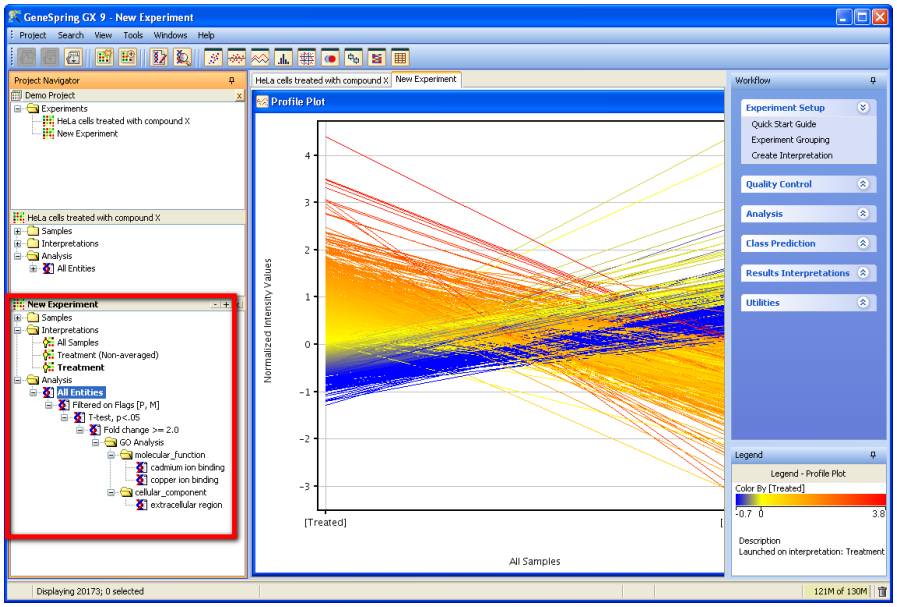
18) Press OK and press Next to move to the next Guided Workflow step; "QC on Samples". This window will show some Quality Metrics from the Feature Extraction Files and shows a PCA plot to determine if the replicates are grouped together. See the manual for more on this, but for the purpose of this demonstration we will step through the windows without any interference, since most of the defaults are good and appropriate.


Guided Workflow - Find Differential Expression (Step 3 of 7)
Steps
1. Summary Report
2. Experiment Grouping
3. QC on samples
4. Filter Probesets
5. Significance Analysis
6. Fold Change
7. GO Analysis
QC on samples
Sample quality can be assessed by examining the values in the PCA plot and other experiment specific quality plots. To remove a sample from your experiment, select the sample from any of the views and click on the Add/Remove button. If a sample is removed, re-summarization of the remaining samples will be performed.
Displaying 6 out of 6 samples retained in the analysis. To change, use the "Add/Remove Samples" button below.
Experiment Grouping
Samples
Treatment
US22502705_251209747382_... Untreated
US22502705_251209747387_... Untreated
US22502705_251209747392_... Treated
US22502705_251209747393_... Treated
US22502705_251209747394_... Untreated
US22502705_251209747404_... Treated
Quality Control Metrics - Report
Experiment Grouping
Quality Control Metrics - Plot
Metric Values
QC Metrics
PCA Scores
PCA Component 2
PCA Component 1
X-Axis
PCA Component 1
Y-Axis
PCA Component 2
Legend - Experiment Grouping
Add/Remove Samples
Help
<< Back
Next >>
Finish
Cancel

19) Press <Next> for the next five screens and read the text at the top to understand what you are looking at. You may want to make some screenshots or save the images you see as JPEGs for your lab notebook. The defaults in all cases are fine and at the last window press the <Finish> button to end the guided workflow.



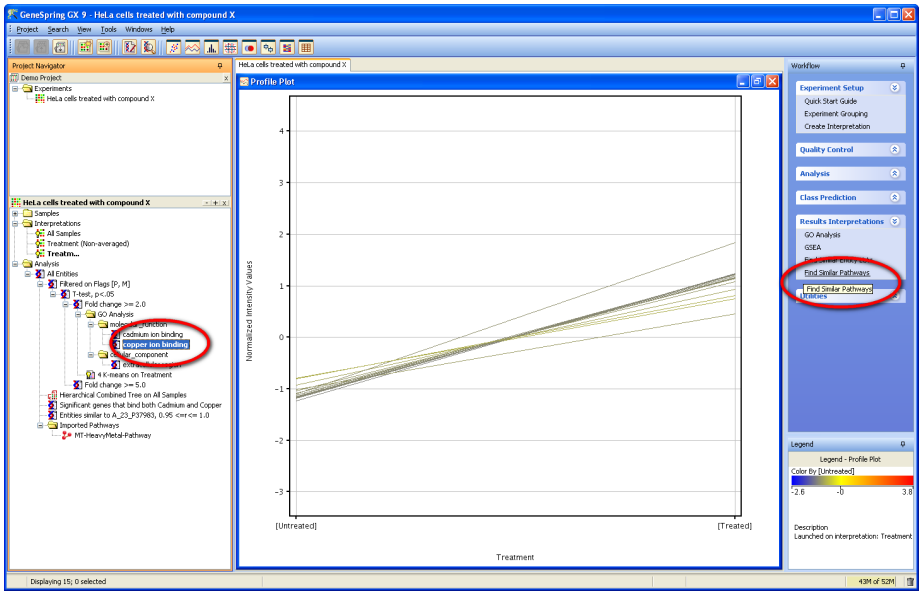
20) After the guided workflows are finished, the main GeneSpring GX window shows all the results of the guided workflow.



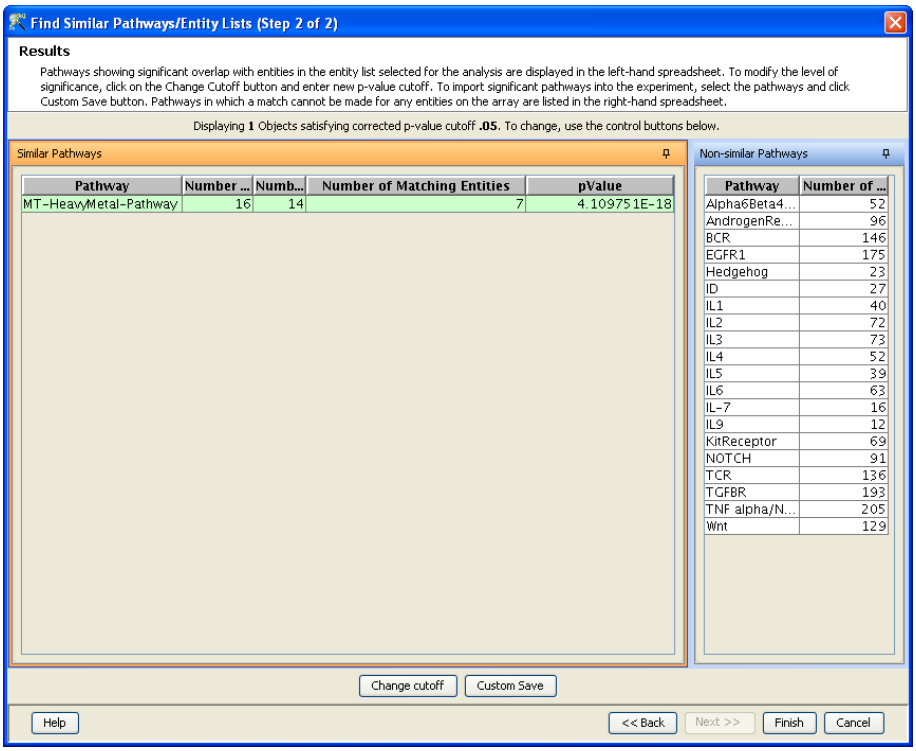
21) The result of the analysis is presented in a hierarchical manner, making it easy to see the steps that were done in the guided workflow and which steps were done.

At this point you have created a new experiment and done some preliminary analysis using the guided workflows. One extra step we invite you to explore is to use the Pathway Analysis functionality in GeneSpring GX 9. We will explore if there is a pathway for which that entity list is significantly enriched.

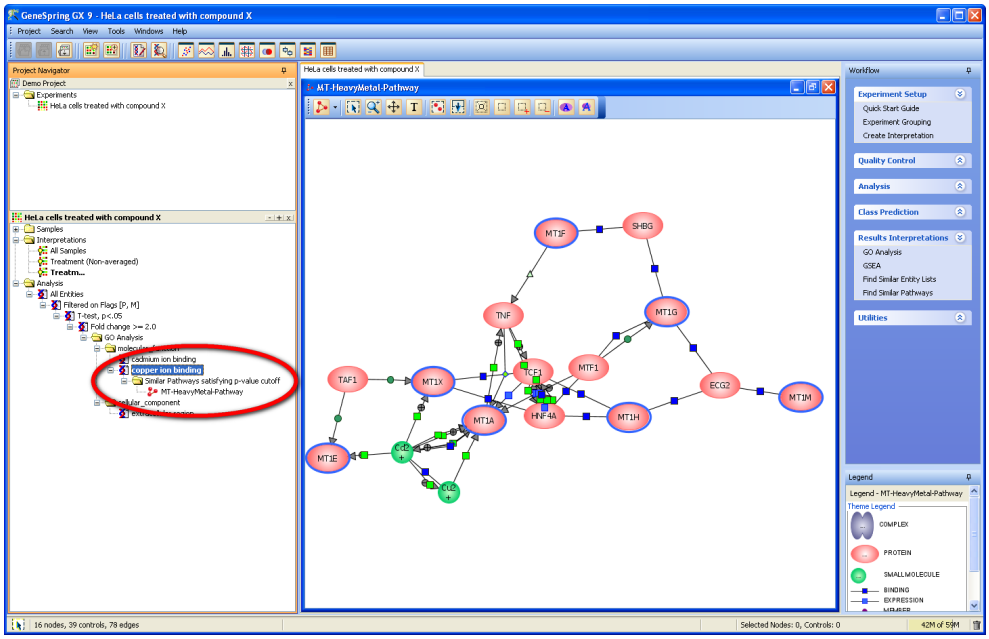
22) Select the Entity List "copper ion binding" by single clicking the list and select the "Find Similar Pathways" function in the Workflow panel on the right, under the "Results Interpretation" section



23) The resulting window will show that that there is one pathway, called "MT-HeavyMetal-Pathway" for which the entity list has 7 of its 16 entities being part of that pathway, showing that there is some definite connection between a large portion of the genes in the list and the P value of 4.1E-18 shows that this finding can be considered to be rather significant


Find Similar Pathways/Entity Lists (Step 2 of 2)
Results
Pathways showing significant overlap with entities in the entity list selected for the analysis are displayed in the left-hand spreadsheet. To modify the level of significance, click on the Change Cutoff button and enter new p-value cutoff. To import significant pathways into the experiment, select the pathways and click Custom Save button. Pathways in which a match cannot be made for any entities on the array are listed in the right-hand spreadsheet.
Displaying 1 Objects satisfying corrected p-value cutoff .05. To change, use the control buttons below.
Similar Pathways
Pathway | Number ... | Numb... | Number of Matching Entities | pValue
MT-HeavyMetal-Pathway | 16 | 14 | 7 | 4.109751E-18
Non-similar Pathways
Pathway | Number of ...
Alpha6Beta4... | 52
AndrogenRe... | 96
BCR | 146
EGFR1 | 175
Hedgehog | 23
ID | 27
IL1 | 40
IL2 | 72
IL3 | 73
IL4 | 52
IL5 | 39
IL6 | 63
IL-7 | 16
IL9 | 12
KitReceptor | 69
NOTCH | 91
TCR | 136
TGFBR | 193
TNF alpha/N... | 205
Wnt | 129
Change cutoff
Custom Save
Help
<< Back
Next >>
Finish
Cancel

24) Press <Finish> to save the pathway as a result in the analysis hierarchy. It will be saved as result under the selected entity list. Double clicking the pathway link will open the network diagram.



At this point we invite you to explore all the functionality of the product. We recommend that you consult the manual (Help menu) and other help is provided on the http://genespring.com website in the form of online tutorials.

How can I get more Pathways?

GeneSpring GX 9 comes pre-loaded with a small set of 20 pathways, courtesy of the **Computational Biology Center** at **Memorial Sloan-Kettering Cancer Center**, the **Gary Bader's lab** at the University of Toronto for the **'Cancer Cell Map'**, the **PandeyLab** at Johns Hopkins University and the **Institute of Bioinformatics** (Bangalore, India). The pathway diagrams were loaded through the BioPAX exchange format, an open platform for the distribution of network and pathway information.

The website http://biopax.org has more information about the BioPAX format and has links to a number of other pathway databases that are provided in the BioPAX format, such as KEGG, BioCyc, NCI Cancer Cell Map and many other pathway databases. Other lists of repositories of BioPAX compatible pathways are provided at the Pathguide site (http://www.pathguide.org/).

NOTE: Import of KEGG pathways in the BioPax format requires non-academic users to obtain a license through the licensor, Pathway Solution, Inc. (pws@kegg.org). Other pathway/networks may require similar license agreements and Agilent Technologies, Inc. cannot be held responsible for unlicensed use of network or pathway data.

Where do I get GSEA compatible Gene Sets?

GSEA (Gene Set Enrichment Analysis), developed by Subramanian, Tamayo, et al. (2005, PNAS 102, 15545-15550) and Mootha, Lindgren, et al. (2003, Nat Genet 34, 267-273) from the BROAD institute, described a fixed set of Gene Sets and these can be used in GeneSpring GX 9 in the GSEA tool. Agilent technologies does not have the re-distribution rights to these gene sets and the user of GeneSpring GX 9 is asked to download the gene sets directly from the BROAD Institute website: http://www.broad.mit.edu/gsea/msigdb/downloads.jsp

Download the file "MSigDB.xml" (At the time of the writing of this manual, the latest version was 2.1 and the file was "msigdb_v2.1.xml"). Unzip the file and use the XML file for import into GeneSpring GX 9 using the "Import BROAD GSEA Gene Sets" in the "Utilities" section.

Where is my data from GeneSpring GX 7?

Due to the completely new data organization in GeneSpring GX 9 when compared to previous versions, data that was saved in GeneSpring GX 7 is not automatically available in version 9, but migration tools are provided. There are two options for data migration, depending on need for the old data:

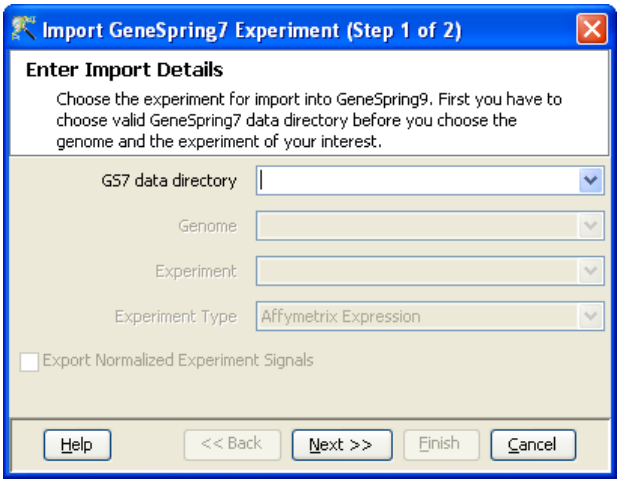
1) Re-create an experiment in GX 9 from the raw files in GX 7
2) Open the normalized data from GX 7 as a new experiment in GX 9

The first option will result in a new experiment in 9, with all the parameter settings etc. maintained, but the normalized values may not be exactly the same. This option is only available for experiments that were created with Agilent FE files and Affymetrix CEL and CHP files.
The second option is available for all experiments from GeneSpring GX 7 and will result in an exact copy of the normalized values, but the samples from which the experiment was created are not available for the creation of a new experiment.
In addition to the experiments, gene lists can also be migrated, but migration of trees and classifications is currently not supported, but these derived objects can easily be re-created in version 9.

Recreating a experiment

For option 1 (re-creation of an experiment), follow the instructions below.

1) Choose menu item "Project -> Import GS7 experiment". Select the GeneSpring GX 7 data directory.

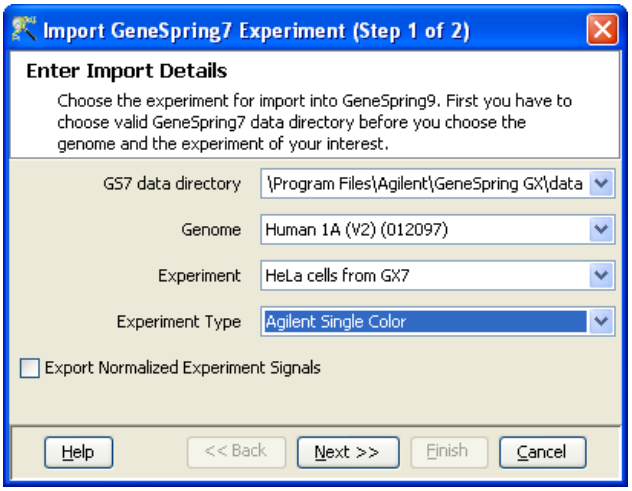

Import GeneSpring7 Experiment (Step 1 of 2)
Enter Import Details
Choose the experiment for import into GeneSpring9. First you have to choose valid GeneSpring7 data directory before you choose the genome and the experiment of your interest.
GS7 data directory
Genome
Experiment
Experiment Type
Affymetrix Expression
Export Normalized Experiment Signals
Help
<< Back
Next >>
Finish
Cancel

2) The default location for the GeneSpring GX data directory is usually:

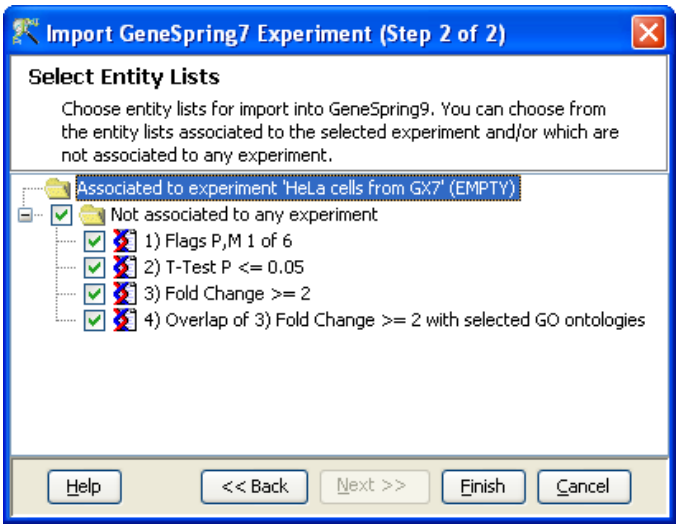
```
C:\Program Files\Agilent\GeneSpring GX\data
```

But may be different in your case

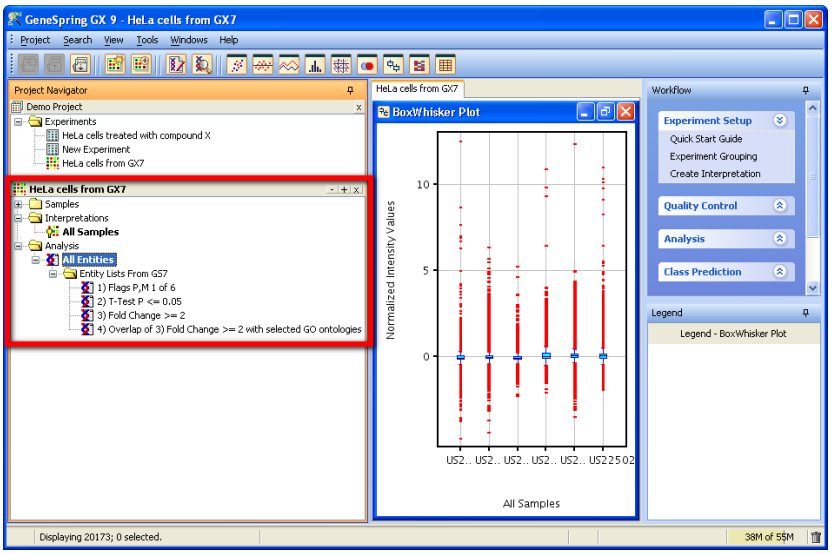
3) After the selection of the data directory, GeneSpring will find all the genomes and experiments that were created in GeneSpring GX 7 and you can select the one experiment you would like to re-create in GX 9


Import GeneSpring7 Experiment (Step 1 of 2)
Enter Import Details
Choose the experiment for import into GeneSpring9. First you have to choose valid GeneSpring7 data directory before you choose the genome and the experiment of your interest.
GS7 data directory
\Program Files\Agilent\GeneSpring GX\data
Genome
Human 1A (V2) (012097)
Experiment
HeLa cells from GX7
Experiment Type
Agilent Single Color
Export Normalized Experiment Signals
Help
<< Back
Next >>
Finish
Cancel

4) Select the appropriate Experiment Type for the Experiment you want to migrate to ensure GeneSpring GX looks for the correct sample attachments and press <Next>
5) The next window will show all the Gene Lists that are present in the genome. GeneSpring GX 9 will attempt to find the gene lists that are were created for the selected experiment, but in many cases this information is not recorded and you can manually choose which gene lists to associate with the experiment, from the folder "Not associated to any experiment". Press <Finish> to continue.


Import GeneSpring7 Experiment (Step 2 of 2)
Select Entity Lists
Choose entity lists for import into GeneSpring9. You can choose from the entity lists associated to the selected experiment and/or which are not associated to any experiment.
Associated to experiment 'HeLa cells from GX7' (EMPTY)
Not associated to any experiment
1) Flags P,M 1 of 6
2) T-Test P <= 0.05
3) Fold Change >= 2
4) Overlap of 3) Fold Change >= 2 with selected GO ontologies
Help
<< Back
Next >>
Finish
Cancel

GeneSpring will now extract the original data files from the samples and start the New Experiment wizard with these samples. All the parameters will be retained, so you can immediately create the interpretations and perform the analysis.

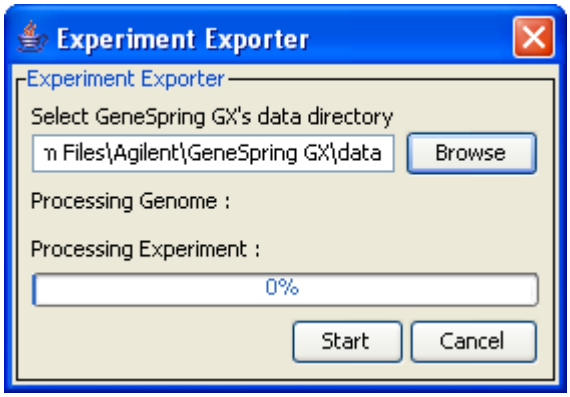


Loading Normalized data as a new experiment

For option 2 (Open normalized data from 7) the follow steps should be followed

1) Run the migration tool once

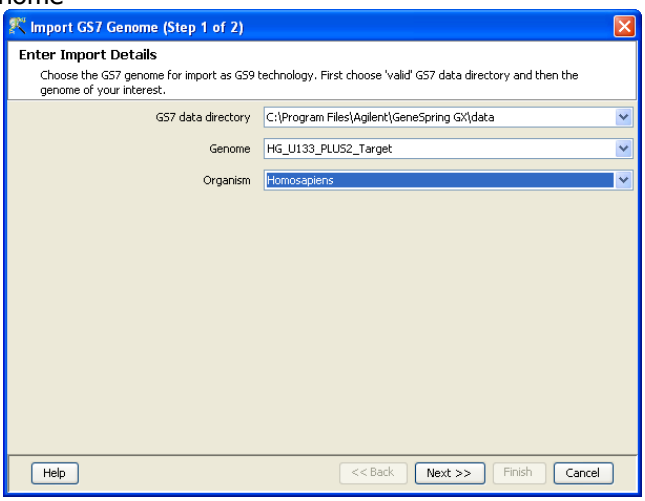
- Menu item "Tools -> Export GS7 Experiments"



- Select the GeneSpring GX 7 data directory and press <Start>.
- This tool only need to be run once and will prepare GX 7 for migration. NOTE: This may take some time if there is a lot of data.

2) Create a custom technology in GeneSpring GX 9

- Menu item "Tools -> Create Custom technology -> Import GS7 genome"



- Select the GeneSpring GX data directory and the genome you would like to migrate
- This tool only needs to be run once for each Genome in GeneSpring GX 7
- This tool will create a library for GX 9 that contains all the annotations for the genes

3) Import a GeneSpring GX 7 experiment

- Menu item "Project -> Import GS7 experiment"
- Choose the option "Export Normalized Experiment Signals"

See the manual for more information about the migration of GeneSpring GX 7 data.